Solution derivation

28 June 2021 17:37

Our Optimization function is,

$$\max_{u_1} \frac{1}{n}\sum_{i=1}^{n} (u_1^T x_i)^2 \quad s.t. \quad u_1^T u = 1$$

This can be written as,

$$Var_1 = \frac{1}{n}\sum_{i=1}^{n} (u_1^T x_i)^2 = \frac{1}{n}\sum_{i=1}^{n} u_1^T x_i x_i^T u_1$$

$$= u_1^T \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \right) u_1$$

$$\boxed{Var = u_1^T S u_1}$$

where $S \rightarrow$ Covariance Matrix of X

Hence optimization problem becomes,

$$\max_{u_1} u_1^T S u_1 \quad s.t. \quad u_1^T u_1 = 1$$

Using Lagrangian Multipliers, we can find the solution to optimization problem.

$$u_1^T S u_1 + \lambda_1 (1 - u_1^T u_1) = 0$$

By diff. partially wrt u_1

$$2 . u_1^T S + (-2) \lambda_1 . u_1^T = 0$$

v

$$2u_1^T S + (-2)\lambda_1 \cdot u_1^T = 0$$

$$\Rightarrow \quad \boxed{S u_1 = \lambda u_1} \longrightarrow (1)$$

By diff. partially wrt λ_1,

$$0 + (1 - u^T u) = 0.$$

$$\boxed{u^T u = 1} \longrightarrow (2)$$

So (1) is the definition of eigen vector.

$$\boxed{A x = \lambda x}$$

A → matrix, x → eigen vector, λ → eigen value

here

$$\boxed{S u_1 = \lambda_1 u_1}$$

S → cov. matrix, u_1 → eigen vector, λ_1 → eigen value.

Hence, By substituting this in variance,

$$V_1 = u_1^T S u_1 = u_1^T \lambda_1 u_1$$

$$= \lambda_1 u_1^T u_1 \qquad \text{since } u_1^T u_1 = 1$$

$$\boxed{V_1 = \lambda_1}$$

i.e. variance of data projected on u_1 is λ_1.

i.e. variance of data projected on u_1 is λ_1.

i.e. our required direction u_1 is eigen vector correspon-
-ding to λ_1.

i.e. u_1 is the first principal component.

This is the solution to variance maximization interpre of PCA

There are many interpretations for PCA.